

**REPORT DOCUMENTATION PAGE**

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information, including suggestions for reducing this burden to Washington, DC 20503.

0350

d  
88  
existing data  
date or any other  
rations and  
(0704-0188).

1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE April 15, 1998	3. REPORT TYPE AND DATES COVERED Final report, 11/1/94 to 11/1/97
4. TITLE AND SUBTITLE  Intermediate Levels of Visual Processing		5. FUNDING NUMBERS  F49620 - 95 - 1 - 0036  2313/BS 61102F	
6. AUTHOR(S)  Ken Nakayama		8. PERFORMING ORGANIZATION REPORT NUMBER  Final Report	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)  Harvard U. 1350 Massachusetts Ave Holyoke Center, 4th Floor Cambridge MA 02138 - 3826		10. SPONSORING / MONITORING AGENCY REPORT NUMBER	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)  AFOSR/NL 110 Duncan Ave Suite B115 Bolling AFB DC 20332-001		11. SUPPLEMENTARY NOTES	
12a. DISTRIBUTION / AVAILABILITY STATEMENT  Approved for public release; distribution unlimited.		12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 Words)  Our progress in trying to understand the psychological properties of human vision can be roughly divided into four main areas. (1) Surface representation. Here we have shown that there is an intermediate level of visual processing, between the analysis of the image and higher order representations related to specific objects. (2) Attention - here we show that attention is not directed to image features but to intermediate level representations (i.e. surfaces). (3) Learning - here we show many forms of perceptual learning, some of which are useful for the directing of focal attention and which are organized at a high level yet which are very machinelike in their operation, others of which are long lasting and seem like insight learning (being all or none), yet which have a representation at a retinotopic level. (4) Faces - here we show that there exist robust representations of faces, these differ from normal face representations in being more efficiently coded and requiring of less attentional resources.			
14. SUBJECT TERMS		15. NUMBER OF PAGES	
		16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT u	18. SECURITY CLASSIFICATION OF THIS PAGE u	19. SECURITY CLASSIFICATION OF ABSTRACT u	20. LIMITATION OF ABSTRACT u

19980430 081

**Final Report Final Report (Oct 1, 1994- Oct 1, 1997)**  
**Grant title: Intermediate levels of visual processing**

Ken Nakayama

**Executive Summary**

Progress in trying to understand the psychological properties of human vision can be roughly divided into four main areas. (1) Surface representation, (2) Attention, (3) Learning (4) Faces. Surface representation and Attention have received the most effort and publications on this topic have been sent to AFOSR. In both cases we argue that low level visual image coding units (analogous to units with known receptive fields) are not adequate to understand the topics at hand, that to understand these topics we need to conceive of representations at the level of mid-level vision. We cannot understand such processing solely at the level of the image (see summaries in Nakayama, He, & Shimojo, 1995; Nakayama and Joseph, 1998).

We have also explored the properties of visual plasticity and show two extremes in possible large continuum of learning. On the one hand, we show a graded unconscious short term memory system for the directing of attention (priming of popout) and facilitating more rapid saccadic eye movements (Maljkovic and Nakayama, 1994, 1996, 1998) Second we show an all or none mechanism in the processing of subjective contours which is surprisingly retinotopic. The great difference between these two memory systems indicates that there are likely to be many intermediate types of learning process, each tailored for different visual functions (Rubin, Nakayama, and Shapley, 1997)

Finally, we have initiated a number of projects on face recognition. Our first study revealed the presence a special class of representations for highly overlearned faces, indicating that these are likely to contain abstract or view-invariant information and also require much less in the way of attentional resources for processing (Tong and Nakayama, in press).

**DTIC QUALITY INSPECTED 3**

### **Papers appearing in print during this period**

1. Nakayama, K. James J. Gibson - An Appreciation. **Psychological Review**, 101, 329-335, 1994.
2. Shimojo, S. and Nakayama, K. Interocularly unpaired zones escape local binocular matching. **Vision Research**, 34, 1875-1881, 1994
3. He, Z.J. and Nakayama, K. Perceived surface shape not features determines correspondence strength in apparent motion. **Vision Research** 34, 2125-2136, 1994
4. Verghese, P. and Nakayama, K. Stimulus discriminability in visual search. **Vision Research**, 34, 2453-2468, 1994
5. Anderson, B.L. and Nakayama, K. Toward a general theory of stereopsis, binocular matching, occluding contours, and fusion **Psychological Review** 101, 414-445, 1994.
6. Maljkovic, V. and Nakayama, K. Priming of popout: I. Role of features, **Memory and Cognition** 22, 657-672, 1994
7. Nakayama, K. and He, Z.J. Attention to surfaces: beyond a Cartesian understanding of visual attention. In Early Vision and Beyond, T. V. Papathomas, ed, M.I.T. Press, Cambridge, 1995.
8. Nakayama, K., He, Z.J. and Shimojo, S. Visual surface representation: a critical link between lower-level and higher level vision. In Kosslyn, S.M. Vision. In Invitation to Cognitive Science. M.I.T. Press, p. 1-70, 1995
9. He, Z.J. and Nakayama, K. Visual attention to surfaces in 3-D space. **Proceedings of the National Academy of Sciences** 92, 11155-11159, 1995.
10. Rubin, N., Nakayama, K., and Shapley, R.M. Enhanced perception of illusory contours in the lower versus the upper visual hemifields. **Science**. 651-653, 1996.
11. Nakayama, K. Binocular visual surface perception. **Proc. Natl. Acad. Sci. USA**, 93, 634-639, 1996.
12. Maljkovic, V. and Nakayama, K. Priming of popout: II. Role of position. **Perception and Psychophysics** 58, 977-991, 1996.
13. Joseph, J.S., Chun, M.M. and Nakayama, K. Attentional requirements in a "preattentive" feature search task. **Nature**, 387, 805-807, 1997

14 Rubin, N., Nakayama, K., and Shapley, R.M. Abrupt learning and retinal size specificity in illusory-contour perception. *Current Biology* 7, 461-467, 1997

### **Manuscripts (in press)**

Joseph, J.S. and Nakayama, K. Memory for the extent of a partially occluded object. (Vision Research, in press).

Nakayama, K. (1997). Vision fin-de-siècle - A reductionistic explanation of perception for the 21st Century? In J. Hochberg & J. E. Cutting (Eds.), Perception and Cognition at Century's End: History, Philosophy, Theory (in press)). Academic Press. Series Handbook of Perception, E. Carterette & M. Friedman (Eds).

Nakayama, K. and Joseph, J. Attention, pattern recognition and popout in visual search. In the Attentive Brain, edited by R. Parasuraman, MIT Press (in press).

Joseph, J.S., Chun, M.M. and Nakayama, K. Reply to Braun (in press, Nature)

Gillam, B. and Nakayama, K. (1998) Quantitative depth for a phantom surface can be based on cyclopean occlusion cues alone. *Vision Res* (in press)

Gillam, B., Blackburn, S. and Nakayama, K. (1998) Unpaired background stereopsis: metrical encoding of depth and slant without matching contours. *Vision Research* (in press)

### **Manuscripts (provisionally accepted and under re-review)**

Maljkovic, V. and Nakayama, K. Priming of Popout III: an example of short term implicit memory *Visual Cognition* ( )

Tse, P., Cavanagh, P. and Nakayama, K. Transformational apparent motion (in Watanabe et al)

McPeck, R.M., Maljkovic, V., and Nakayama, K. Saccades require focal attention and are facilitated by a short-term memory system. *Vision Res* ( )

Tong, F. and Nakayama, K. Robust Representations for faces: evidence from visual search. *J. Exp Psychol. Human Percp and Perf* ( )

### **Published abstracts**

Joseph, J.S., Chun, M.M., Nakayama, K. (1996). Attention plays a role in the perception of three-dimensional structure in shaded cube stimuli.

Investigative Ophthalmology & Visual Science, 37, p S213 (Abstract 975).

Joseph, J. S., Chun, M. M., & Nakayama, K. (1996, November). Attentional requirements in a "preattentive" feature search task. Poster to be presented at the annual meeting of the Psychonomic Society, Chicago, IL.

McPeck, R. M. and Nakayama, K. (1995). Linkage of attention and saccades in a visual search task. *Investigative Ophthalmology and Visual Science (Suppl.)*, 36(4), S354.

McPeck, R. M. and Nakayama, K. (1995). Repetition of Target Color Affects Saccadic Latency and Accuracy. 8th European Conference on Eye Movements, Derby, UK.

Rubin, N., Nakayama, K., Shapley, R. and Grossetete, A. (1996), Abrupt learning in illusory contour perception. *Invest. Ophth. and Vis. Sci. Suppl.* 37, 859.

Rubin, N., Grossetete, A., Nakayama, K. and Shapley, R. (1996), Abrupt transitions in illusory contour perception triggered by specific visual stimuli. *Perception* 25 (Suppl.).

Tong, F.H., Nakayama, K. (1996). Finding your face in a crowd: familiarity and inversion effects in visual search. Poster ARVO 1366, page S297

Maljkovic, V., Nakayama, K. (1996). Brief duration of pop-out priming. ARVO talk, 58, page S14

Tse, P., Nakayama, K. and Cavanagh, P. (1996). The roles of attention in shape change apparent motion. Arvo talk, 977, page S213

Mc Peek, R.M., Skavenski, A., and Nakayama, K., (1996). Saccades in a visual search task show concurrent programming. Arvo talk, 2138 page S471.

Yilmaz, E. and Nakayama, K. Fluctuation of attentional levels during driving *Invest. Ophthal. and Visual Sci.* 1995 Supplement, #4330

Gillam B., Nakayama K., Blackburn S. Subjective contours produced by end-cutting may not appear at the end-cuts. *Invest. Ophthal. and Visual Sci.* 1997, #989.

## **Staff involved during the project period**

Barton Anderson, Post doctoral student  
Elizabeth Bennett, Staff Assistant  
Yue Chen, Post-doctoral fellow  
Raynald Comtois, Software engineer  
Barbara Gillam, visiting scientist  
Anne Grossetete, Research Assistant  
Zijiang He, Ph.D. Post-doctoral fellow  
Julian Joseph, Post-doctoral fellow  
Vera Maljkovic, Pre-doctoral student  
Robert McPeck Pre-doctoral student  
Nava Rubin, Ph.D Post-doctoral fellow  
Frank Tong, Pre-doctoral student  
Peter Tse, Pre-doctoral student  
Emre Yilmaz, Pre-doctoral student

## **Ph.D Thesis Completed**

Robert McPeck. Attention and eye movements in visual search. May 1997

## **Summary of Work Conducted**

**1. Reviews/Position Papers** - With many papers appearing in diverse journals, I felt it useful to write synthetic/theoretical papers to integrate the many studies into coherent accounts. In this regard 4 such papers have either been published or completed during the current grant period.

Nakayama, He, and Shimojo, (1995) summarized and advocated the main theoretical points of our work on surface representation.

--

Nakayama and Joseph (1997) is a theoretical chapter, making the case against the view (Treisman, 1985; Julesz, 1984) that built in primitives (akin to receptive field or feature analyzers) determine the properties of visual search.

Nakayama (1997) is a historical survey of vision over the past hundred years, arguing for a broadly eclectic approach advocated in this proposal.

Nakayama (1994) was a panegyric for J.J. Gibson, explaining how his work antedated current neurophysiological studies on the perception of motion and indicating Gibson's possible future influence.

**2. Priming of Popout (PoP)** Bravo and Nakayama (1992) devised a visual search task which required that the subject use focal attention to identify the shape of a target which had the odd feature. In so doing, they found a characteristic signature of focal attention for such popout tasks. When target and distractor color remained the same from trial to trial (say a red target among green distractors), reaction times were flat with increasing distractor number. When target and distractor color could vary randomly from trial to trial (target and distractor colors could interchange), reaction times were overall much longer and there was a characteristic *decrease* in reaction time as distractor number was increased. In a series of studies conducted by Maljkovic and Nakayama (1994;1996,1997) we established that it was the presence of priming that aided in the deployment of attention to the odd target which was responsible for the better performance in the blocked trials case. The priming was very characteristic showing the priming for the feature identity and position of the odd target but not its shape. For example, priming was restricted to the color which previous attracted spatial attention and to the place that attention was allocated. It did not show any evidence for priming for the dimension for which attention was required, i.e. shape. Priming lasted for approximately 5-8 trials, spanning a range of approximately 30 seconds with the intertrial intervals chosen. We also established that conscious awareness of the color of the upcoming trial was irrelevant indicating that the priming was more or less unconscious, not subject to observer control. Finally in the most recent paper in the series (Maljkovic and Nakayama, 1997) we established that PoP represents a new class of memory phenomena, short term implicit memory. (note that all previous accounts of implicit memory have been long-term). It is a almost machine-like process that effectively reduces the time to deploy attention in situations that recur in the moment.

**3. Saccades require focal attention and show short term implicit priming.** In the usual visual search tasks, we find the characteristic flat reaction time function as the number of distractors is increased. Bravo and Nakayama, requiring focal attention, found a very different function when distractor number was increased, *reaction times fell*. As mentioned above, however, this occurred only in the mixed as opposed to the blocked trials case. In a recent series of studies of the attentional requirements of saccadic eye movements, we have shown that Bravo and Nakayama's characteristic signature for focal attention in popout tasks is also mirrored in the pattern of saccadic latencies (see accompanying figure). This indicates that the deployment of attention to a saccadic target requires focal as opposed to distributed attention. This was further confirmed by showing that the same memory kernel analysis so successfully used in studies of attention applied to saccadic eye movement latencies as well. These results taken together, indicate a very strong connection between eye movements and the deployment of *focal* and not distributed attention (McPeck, Maljkovic, and Nakayama, 1997).

McPeck, Skavenski and Nakayama (1996) have also made measurements on corrective saccades in visual search tasks, showing that very short latency corrections are pre-programmed. Thus corrective eye movements occurring between 0 and 80 milliseconds of the end of a first saccade do not interrupt the first saccade (as indicated

from a "phase plane" analysis of the dynamics of the velocity trajectory) but are separate and concurrently planned eye movements. The existence of such concurrently planned saccades is incompatible with current neural models of eye movement control. The writing of the second part of this study is nearing completion and will be submitted to the Journal of Neurophysiology.

#### **4. Unpaired points in stereoscopic vision**

In Shimojo and Nakayama (1994) and Anderson and Nakayama (1994), we dealt with the issue of DaVinci stereopsis, showing in specific ways how unpaired points do not make matches in occlusion zones (Shimojo and Nakayama, 1994) yet are likely to define bounding contours and to influence binocular matching (Anderson and Nakayama, 1994). More recently, in Gillam and Nakayama (1998a) we have shown that unpaired background regions are also correctly perceived (are not seen as rivalrous) and also serve to determine the depth of otherwise ambiguous regions in a way that is in line with most probable scenes that could have given rise to them. In the accompanying diagram, we show a stereoscopic figure from this paper. Of importance to note is that the unpaired background region can be very far from paired regions of defined disparity. This contradicts Nakayama and Shimojo's (1990) conjecture that unpaired points operate only locally and are thus a variant of conventional stereopsis. Similarly, it cannot be explained by the mechanistic receptive field model outlined in Anderson and Nakayama (1994) because the distances between unpaired regions and paired regions is too large. As such, it indicates that we must think of more global mechanisms to bridge the gap across space to link unpaired to paired regions to obtain appropriate depth. What is clear is that we need to understand how such delocalized mechanisms can maintain eye of origin identity over such long distances.

**5. Stereoscopic Subjective contours support a surface as opposed to image interpretation of visual search.** One of the more recurring notions in the subjective contour literature has been the role of line endings and abutting lines. Most theories have suggested that subjective contours are preferentially formed orthogonal to such line endings. Earlier I indicated a counter example to this, citing the work of Gillam (1987) where randomly oriented lines led to even more vivid subjective contours. Here we also argue against the line abutment idea by developing stereograms which have much weaker line abutments yet stereoscopic subjective contours are very strong. All of these findings only make sense when we regard the formation of subjective contours as involving the explicit interpretation of surfaces. The stereogram below, we show that when the bottom sticks are seen in front, we have a strong sense of a subjective contour such that our interpretation is consistent with the 3-D interpretation of the scene depicted in the inset on the far right. This is not seen when the lower sticks are seen in back. This indicates that one obtains occluding contours preferentially from a planar surface but that significant extensions of this surface are allowed such that they project in front of the occluded surface. Thus, the visibility of the sticks protruding from the occluding surface does not destroy the subjective



contour as might be expected from any extant theory of subjective contour formation. (Gillam and Nakayama, 1997)

#### **6. Visual attention is deployed to surfaces not features**

Although we now recognize that attention plays a significant role in vision, its spatial deployment and spread in the third dimension has not been well understood. Previous thinking has been wedded to the idea that attention is the modulation of particular neurons with characteristic receptive fields. Thus, visual attention was regarded as attending to features. Nakayama and Silverman (1986) argued that attention could be deployed to particular values of binocular disparity. Thus, I still accepted the then current thinking regarding the importance of features in the deployment of attention, suggesting that a class of say near or far cells (Poggio, 1989) could be selectively modulated. In the current work, we find that this idea is unlikely to be correct, that the selective process of attention is primarily to higher order representations, not elementary image dimensions.

In visual search experiments we show that we cannot easily focus attention across iso-disparity or iso-depth loci unless they are part of a well-formed surface with locally co-planar elements. In the accompanying figure, depict this situation by indicating that the search array shown in (a) (where all the elements are coplanar) is easy to search for the odd colored middle disparity target. This cannot be done with ease for case (b) where there the search array is also at the same stereoscopic disparity but search is very difficult. Nevertheless, we can easily spread our attention selectively across well-formed surfaces which span an extreme range of stereoscopic depths as long as the elements are co-planar (not shown here). We also asked the complementary question. Is the ease over which attention can spread over surfaces something automatic or is it something that is under voluntary control. Our next set of experiments suggested that the spread of attention is more or less involuntary. We examined this by seeing whether subjects could restrict attention to within a defined section of a large surface as easily as restricting it to a region of comparable size which were on a separate isolated surfaces. In cueing experiments, we found that attention could not be so effectively sequestered within a surface as between two separate surfaces, indicating that the spread of attention across surfaces has an obligatory involuntary component. It is likely to be related to the ease by which it spreads across surfaces. We conclude that attention cannot be efficiently allocated to arbitrary depths and extents in space but is linked to and spreads automatically across perceived surfaces (see He and Nakayama, PNAS, 1995).

**7. Up-down asymmetry in illusory contour formation:** ( We have shown that subjective contours are more salient in the lower as opposed to the upper hemi-field and that objective tasks based on the emergence of subjective contours are performed better in the lower as opposed to the upper field. We suggest the following broad interpretation of this result. 1). Richly connected areas are more likely to form brain specialized brain modules 2) the upper and lower field representation of area V2 (and higher) are separated due to the topographic breakage of the mapping of retina to extra-

striate cortex. Such breakage does not exist at V1 or earlier. The breakage of upper and lower field representation in extra-striate cortex provides the opportunity an emergence of specialized processing for the upper and lower visual field. If these ideas are valid, then it is of interest to see whether other sorts of specializations are also apparent, showing upper vs lower field differences. Should this be the case, it could reflect differences in visual processing after the primary striate cortex. (Rubin, Shapley and Nakayama, Science, 1995)

#### **8. Perceptual learning can be very abrupt and stimulus specific suggesting that one cannot regard it as a passive Hebbian process.**

Recently there has been a revival of interest in perceptual learning. The presumed mechanistic underpinnings of this learning have generally been some variant of incremental learning, such as Hebbian association. Such work connects psychological studies of perceptual learning to presumed changes of synaptic efficacy in the visual cortex. As such, the slowness of the learning, requiring hundreds of trials over days is consistent with the slower changes in synaptic modifiability seen in physiological studies. In studying such forms of learning, there has been the implicit yet strong presumption that it is a passive process, simply governed by incremental laws of association and likely to be modeled as something similar to Hebbian learning. This is to be contrasted with very different forms of learning, say insight learning, where the transitions from one state to another are abrupt and presumably reflect higher order understanding. The existence of insight learning while acknowledged is often not studied scientifically for several reasons. First, it has been very hard to document in the laboratory with any degree of rigor or repeatability. Second, the mechanism of insight learning is not easily handled by known physiological mechanisms as is gradual incremental learning. As such, questions as the mechanism of insight learning has generally been ignored by physiologists or at least considered as a very separate and higher order process, sufficiently different to put it into a very different category (i.e. safe to be ignored). That perceptual learning which has been deemed to reveal changes in retinotopic areas might be sudden has not been considered.

Our results are interesting on two counts. First, we have developed a paradigm where very abrupt learning can be brought under laboratory control and measured quantitatively. Second, we have found that abrupt learning is not as separable from early incremental learning as one might suppose because it shows one of the presumed "telltale" signatures of passive low-level incremental learning, i.e. namely its retinotopicity. (Rubin, Shapley, and Nakayama, 1997)

#### **9. Attentional requirements for so called "popout tasks"**

Most studies of visual attention have made the tacit assumption that certain visual tasks do not require attention. I have been arguing against this idea for over six years (Nakayama, 1990) and during the grant period, I wrote another more comprehensive review essay which should receive wider circulation and I trust should receive a more sympathetic audience (see Nakayama and Joseph, 1997). Basically and bluntly, my theoretical point is that much of the theory on attention, particularly having to do with

easy visual search (i.e. the pre-attentive/attentive distinction made by Julesz, Treisman and others) is simply wrong. Empirically and to provide further support for this view, we have provided strong evidence that those tasks generally thought not to require attention, namely "popout" tasks really do (Joseph, Chun, and Nakayama, 1996). Our point is that researchers have been satisfied with weak manipulations of attention and then when they are made much stronger, its importance becomes dominant. (Joseph, Chun, and Nakayama, Nature, 1997, 1998).

**10. Visual processing of the human face.** We have also initiated a long term effort to develop a large and very flexible laboratory facility to examine the perception of the human face. So far-

(1) We have identified what we call robust representations of faces, a representation going far beyond fluent familiarity which indicates a more efficient view-invariant coding process requiring much less attentional resources for identification. (Tong and Nakayama, J. exp Psychol., under revision)